

WHAT IS CLAIMED IS:

1. A computer storage system, comprising:
a plurality of disk drives for storing parity groups, each parity group comprising storage blocks, said storage blocks comprising one or more data blocks and a parity block associated with said one or more data blocks, each of said storage blocks stored on a separate disk drive such that no two storage blocks from a given parity group reside on the same disk drive;
a redistribution module to dynamically redistribute parity groups by combining some parity groups to improve storage efficiency.
2. The computer storage system of Claim 1, wherein a size of a first parity group is larger than a size of a second parity group.
3. The computer storage system of Claim 1, further comprising metadata to specify which disk drive in said plurality of disk drives contains a desired storage block.
4. The computer storage system of Claim 1, further comprising metadata that specifies at least one Gnid-string.
5. The computer storage system of Claim 4, wherein a one-to-one correspondence exists between said at least one Gnid-string and a directory.
6. The computer storage system of Claim 4, wherein said Gnid-string comprises a collection of gnids.
7. The computer storage system of Claim 6, wherein each of said gnids comprises information for locating a specified gnode.
8. The computer storage system of Claim 7, wherein said information for locating a specified gnode comprises a pointer to said specified gnode.
9. The computer storage system of Claim 7, wherein said specified gnode comprises file attributes for a file corresponding to said gnode.
10. The computer storage system of Claim 11, wherein said file attributes include at least one of a file id, a file access time, a file creation time, and a file modification time.
11. The computer storage system of Claim 7, wherein said specified gnode comprises information for locating a first gee of a plurality of gees corresponding to said gnode.

12. The computer storage system of Claim 11, wherein said plurality of gees comprises gnode gees and data gees.

13. The computer storage system of Claim 12, wherein each of said gnode gees comprises information to specify an extent.

5 14. The computer storage system of Claim 12, wherein each of said data gees comprises information to specify a first logical disk block and information to specify which of said disk drives contains said first logical block.

15. The computer storage system of Claim 12, wherein said plurality of gees further comprises parity gees.

10 16. The computer storage system of Claim 15, wherein each of said parity gees comprises information regarding location of parity data for one or more preceding data gees in said plurality of gees.

17. The computer storage system of Claim 11, wherein a parity group comprises a first set of one or more data gees and an associated parity gee.

15 18. The computer storage system of Claim 1, wherein each data gee identifies a block of data and said parity gee identifies a parity block.

19. The computer storage system of Claim 1, wherein said disk drives are provided to a Fibre channel network.

20 20. The computer storage system of Claim 1, wherein a file is comprised of one or more of said parity groups.

21. The computer storage system of Claim 1, further comprising metadata to specify a selected logical disk block and to specify which of said disks that contains said selected logical block.

25 22. The computer storage of Claim 1, wherein a size of a first parity group is larger than a size of a second parity group of the same file.

30 23. The computer storage system of Claim 1, wherein said redistribution module combines a first parity group having m data blocks and one parity block and a second parity group having n data blocks and one parity block to produce a combined parity group having $m+n$ data blocks and one parity block, where $m+n$ is at most one less than the number of said disk drives.

24. The computer storage system of Claim 23, wherein said redistribution module run in a background process during file system operations.

25. The computer storage system of Claim 19, further comprising an allocation module to split a first parity group having k data blocks and one parity block into a second parity group having m data blocks and one parity block and a third parity group having n data blocks and one parity block, where $k \geq m+n$.

26. The computer storage system of Claim 25, wherein said allocation module runs in a background process during file system operations.

27. The computer storage system of Claim 25, wherein said first parity group is an unused parity group.

28. A method for storing data in a computer network, comprising:
determining a size of a parity group in response to a write request, said size describing a number of data blocks in said parity group;
arranging at least a portion of data from said write request according to said data blocks;

computing a parity block for said parity group;
storing each of said data blocks on a separate disk drive such that no two data blocks from said parity group reside on the same disk drive;
storing said parity block on a separate disk drive that does not contain any of said data blocks; and

redistributing said parity groups to improve storage efficiency.

29. The method of Claim 28, further comprising storing metadata to describe a disk and logical block location of each of said data blocks and said parity block.

30. The method of Claim 28, wherein said redistributing comprises combining a first parity group having a first size and a second parity group having a second size to produce a combined parity group having a third size, wherein said third size specifies a number of blocks that is, at most, one less than the number of disk drives available to store data from said parity group.

31. The method of Claim 28, wherein said redistributing comprises splitting a first parity group into a second parity group and a third parity group.

32. The method of Claim 28, further comprising allocating a new parity group from a pool of available parity groups.

33. The method of Claim 32, further comprising generating parity groups for said pool of available parity groups from unused disk space.

5 34. The method of Claim 32, further comprising generating a plurality of differently-sized parity groups for said pool of available parity groups from unused disk space.

10 35. The method of Claim 34, further comprising splitting a parity group in said pool of parity groups to produce two smaller parity groups in said pool of available parity groups.

15 36. The method of Claim 34, further comprising combining one or more parity groups in said pool of parity groups to produce a combined parity group in said pool of available parity groups, said combined group larger than said one or more parity groups.